

```

(5')  1  SDLEQERRAKEKLQEQQ
      18  SDLEQDRLAKEKLQEQQ
      35  SDLEQERLAKEKLQEQQ
      52  SDLEQERRAKEKLQEQQ
      69  SDLEQERRAKEKLQEQQ
      86  SDLEQDRLAKEKLQEQQ
     103  SDLEQERRAKEKLQEQQ
     120  SDLEQERRAKEKLQEQQ
     137  SDLEQERLAKEKLQEQQ
     154  SDLEQERRAKEKLQEQQ
     171  SDLEQERRAKEKLQEQQ
     188  SDLEQERRAKEKLQEQQ
     205  RDLEQ

     210  RKADTKKNLERKKKEHGDILAEDLYGRLEIP
     240  AIELPSENERGYYPHQSSLPQDNRGNSRD
     270  SKEISIIIEKTNRESITTNVEGRRDIHKGHL
     300  EEKKDGSIKPEQKEDKS      316 (3')

```

FIGURE 1

(5') 1 AAAGCGATCTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
52 AAAGCGATTTAGAACAAGATAGACTTGCTAAAGAAAAGTTACAAGAGCAGC
103 AAAGCGATTTAGAACAAGAGAGACCTTGCTAAAGAAAAGTTGCAAGAACAAC
154 AAAGCGATCTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
205 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
256 AAAGCGATTTAGAACAAGATAGACTTGCTAAAGAAAAGTTACAAGAGCAGC
307 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
358 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
409 AAAGCGATTTAGAACAAGAGAGACCTTGCTAAAGAAAAGTTGCAAGAACAAC
460 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
511 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
562 AAAGCGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAGCAGC
613 AAAGAGATTTAGAACA
630 AGGAAGGCTGATACGAAAAAAATTTAGAAAGAAAAAGGAACATGGAGAT
681 ATATTAGCAGAGGATTTATATGGTCGTTTAGAAAATACCAGCTATAGAACTT
732 CCATCAGAAAAATGAACGTGGATATTATATACCACATCAATCTTCTTACCT
783 CAGGACAACAGAGGGAATAGTAGAGATTCGAAGGAAATATCTATAATAGAA
834 AAAACAAATAGAGAATCTATTACAACAAATGTTGAAGGACGAAGGGATATA
885 CATAAAGGACATCTTGAAGAAAAGAAAGATGGTTCAATAAAACCAGAACAA
936 AAAGAAGATAAATCT 950 (3')

FIGURE 2

RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVJSVQQEQQ
HNVEEKVEESVEENDEESVEENVEENVEENDDDGSVASSVEESI
ASSVDESIDSSIEENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVE
ESVAPSVEESVAEMLKER

FIGURE 3

5' GAA TTC CGT GAT GAA CTT TTT AAT GAA TTA TTA AAT AGT GTA GAT
 GTT AAT GGA GAA GTA AAA GAA AAT ATT TTG GAG GAA AGT CAA GTT AAT
 GAG GAT ATT TTT AAT AGT TTA GTA AAA AGT GTT CAA CAA GAA CAA CAA
 CAC AAT GTT GAA GAA AA AGT TGA AGA AAG TGT AGA AGA AA ATG ACG
 AAG AAA GTG TAG AAG AAA ATG TAG AAG AAA ATG TAG AAG AAA ATG
 ACG ACG GAA GTG TAG CCT CAA GTG TTG AAG AAA GTA TAG CTT CAA GTG
 TTG ATG AAA GTA TAG ATT CAA GTA TTG AAG AAA ATG TAG CTC CAA CTG
 TTG AAG AAA TCG TAG CTC CAA CTG TTG AAG AAA TTG TAG CTC CAA GTG
 TTG TAG AAA AGT GTG CTC CAA GTG TTG AAG AAA GTG TAG CTC CAA GTG
 TTG AAG AAA GTG TAG CTG AAA TGT TGA AGG AAA GGA ATT C 3'

FIGURE 4

<u>DELFNELLNSVDVNGEVKENILEESQ</u>	<u>NSRDSKEISIIIEKTNRESITTNVEGRDDIHK</u>	<u>LSA-TER</u>
<u>LEESQVNDDDIFSNLSLVKSVOQQEQHHNV</u>		<u>729S-NRI</u>
<u>VEKCAPSVEESVAPSVEESVAEMLKER</u>		<u>729S-NRII</u>
		<u>729S-Rep</u>

FIGURE 5

NUCLEOTIDE SEQUENCE OF THE LSA GENE
5' END

(NON-CODING 5' END)

1 AAAGTATACATCTTCCTTCTTTACTTCTTAAA

(CODING 5' END, UNIQUE)

33 ATGAAACATATTTTGTACATATCATTTTACTTTATCCTTGTTAATTTATTG
84 ATATTTTCATATAAATGGAAAGATAATAAAGAATTCTGAAAAAGATGAAATCA
135 TAAAATCTAACTTGAGAAGTGGTTCTTCAAATTCTAGGAATCGAATAAATGA
186 GGAAAATCACGAGAAGAAACACGTTTTATCTCATAATTCATATGAGAAAAC
237 AAAAATAATGAAAATAATAAATTTTTTCGATAAGGATAAAGAGTTAACGATGT
288 CTAATGTAAAAAATGTGTCACAAACAAATTTCAAAGTCTTTTAAGAAATCT
339 TGGTGTTCAGAGAATATATTCCTTAAAGAAAATAAATTAAATAAGGAAGGG
390 AAATTAATTGAACACATAATAAATGATGATGACGATAAAAAAAAAATATATTA
441 AAGGGCAAGACGAAAACAGACAAGAAGATCTTGAAGAAAAAGCA

(CODING 5' END, repetitive)

492 GCTAAAGAAAAGTTACAGGGGCAACAAAGCGATTTCAGAACAAGAGAGACGT
543 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACTT
594 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACGT
645 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACTT
696 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACGT
747 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACGT
798 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACTT
849 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTTAGAACAAGATAGACTT
900 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTTAGAACAAGAGAGACGT
951 GCTAAAGAAAGGTTGCAAGAACAACAAAGCGATTTTAGA 988

FIGURE 6

DNA sequence 956 b.p. ATGAAACATATT ... AAGCGATTTAGA linear

FIGURE 7A

781 / 261	811 / 271	
CAA GAA CAA CAA AGC GAT TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG		
gln gln gln ser asp leu gln gln gln gln ala lys glu lys leu gln gln gln		
841 / 281	871 / 291	
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA CAA AGC GAT		
gln ser asp leu glu gln asp arg leu ala lys glu lys leu gln glu gln gln ser asp		
901 / 301	931 / 311	
TTA GAA CAA GAG AGA CGT GCT AAA GAA AGG TTG CAA GAA CAA CAA AGC GAT TTA		
leu glu gln glu arg arg ala lys glu arg leu gln glu gln gln ser asp leu		

FIGURE 7C

NUCLEOTIDE SEQUENCE OF THE LSA GENE
3' END

(CODING 3' END, REPETITIVE)

```
1  CAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
37 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
88 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACTT
139 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
190 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
241 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
292 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACGT
343 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
394 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
445 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
496 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
547 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
598 GCTAAAGAAAAGTTGCAAGAGCAGCAAAGAGATTTAGAACAA
```

(CODING 3' END, UNIQUE)

```
640 AGGAAGGCTGATACGAAAAAAATTTAGAAAGAAAAAAGGAACATGGAGAT
691 ATATTAGCAGAGGATTTATATGGTCGTTTAGAAATACCAGCTATAGAACTT
742 CCATCAGAAAATGAACGTGGATATTATATACCACATCAATCTTCTTTACCT
793 CAGGACAACAGAGGGAATAGTAGAGATTCCAAGGAAATATCTATAATAGAA
844 AAAACAAATAGAGAATCTATTACAACAAATGTTGAAGGACGAAGGGATATA
895 CATAAAGGACATCTTGAAGAAAAGAAAGATGGTTCAATAAAACCAGAACAA
946 AAAGAAGATAAATCTGCTGACATACAAAATCATACATTAGAGACAGTAAAT
997 ATTTCTGATGTTAATGATTTTCAAATAAGTAAGTATGAGGATGAAATAAGT
1048 GCTGAATATGACGATTCATTAATAGATGAAGAAGAAGATGATGAAGACT
1099 TAGACGAATTTAAGCCTATTGTGCAATATGACAATTTCCAAGATGAAGAAA
1150 ACATAGGAATTTATAAAGAACTAGAAGATTTGATAGAGAAAAATGAAATT
1201 TAGATGATTTAGATGAAGGAATAGAAAAATCATCAGAAGAATTATCTGAAG
1252 AAAAAATAAAAAAAGGAAAGAAATATGAAAAACAAAGGATAATAATTTTA
1303 AACCAAATGATAAAAGTTTGTATGATGAGCATATTAATAAAATATAAAATG
1354 ATAAGCAGGTTAATAAGGAAAAGGAAAAATTCATAAAATCATTGTTTCATA
1405 TATTTGACGGAGACAATGAAATTTTACAGATCGTGGATGAGTTATCTGAAG
1456 ATATAACTAAATATTTTATGAAACTATAA (stop)
```

(NON-CODING 3' END)

1485 AAGGTTATATATTT 1498

FIGURE 8

LSA.3'.ALL -> 1-phase Translation

DNA sequence 1496 b.p. CAAGAACAACAA ... GGTATATATTT linear

1 / 1	31 / 11	
CAA GAA CAA CAA AGC GAT CTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA		
gln glu gln gln ser asp leu glu gln glu arg ala lys glu lys leu gln glu gln		
61 / 21	91 / 31	
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT		
gln ser asp leu glu gln asp arg leu ala lys glu lys leu gln glu gln ser asp		
121 / 41	151 / 51	
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT CTA GAA CAA		
leu glu gln glu arg leu ala lys glu lys leu gln glu gln ser asp leu glu gln		
181 / 61	211 / 71	
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAG AGA CGT		
glu arg arg ala lys glu lys leu gln glu gln ser asp leu glu gln glu arg arg		
241 / 81	271 / 91	
GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA		
ala lys glu lys leu gln glu gln ser asp leu glu gln asp arg leu ala lys glu		
301 / 101	331 / 111	
AAG TTA CAA GAG CAG CAA AGC GAT TTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA		
lys leu gln glu gln ser asp leu glu gln glu arg arg ala lys glu lys leu gln		

FIGURE 9A

841 /	281	871 /	291
GAA AAA	ACA AAT AGA GAA	TCT ATT ACA	ACA AAT GTT GAA GGA CGA AGG GAT ATA CAT AAA
glu lys	thr asn arg glu	ser ile thr	thr asn val glu gly arg asp ile his lys
901 /	301	931 /	311
GGA CAT	CTT GAA AAG AAA	GAT GGT TCA ATA AAA	CCA GAA CAA AAA GAA GAT AAA TCT
gly his	leu glu lys	lys asp gly ser ile	lys pro glu gln lys glu asp lys ser
961 /	321	991 /	331
GCT GAC	ATA CAA AAT CAT	ACA TTA GAG ACA GTA AAT ATT TCT GAT	GTT AAT GAT TTT CAA
ala asp	ile gln asn his	thr leu glu thr val	asn ile ser asp val asn asp phe gln
1021 /	341	1051 /	351
ATA AGT	AAG TAT GAG GAT	GAA ATA AGT GCT GAA TAT GAC GAT	TCA TTA ATA GAT GAA GAA
ile ser	lys tyr glu asp	glu ile ser ala glu tyr asp asp	ser leu ile asp glu glu
1081 /	361	1111 /	371
GAA GAT	GAT GAA GAC TTA	GAC GAA TTT AAG CCT ATT GTG CAA	TAT GAC AAT TTC CAA GAT
glu asp	asp glu asp leu	asp glu phe lys pro ile val gln tyr asp	asn phe gln asp
1141 /	381	1171 /	391
GAA GAA	AAC ATA GGA ATT	TAT AAA GAA CTA GAA GAT TTG ATA GAG	AAA AAT GAA AAT TTA
glu glu	asn ile gly ile	tyr lys glu leu glu asp leu ile	glu lys asn glu asn leu
1201 /	401	1231 /	411
GAT GAT	TTA GAT GAA GGA	ATA GAA AAA TCA TCA GAA GAA TTA TCT	GAA AAA ATA AAA
asp asp	leu asp glu gly	ile glu lys ser ser glu glu	ser glu lys ile lys

FIGURE 9C

1261 / 421	1291 / 431	
AAA GGA AAG AAA TAT GAA AAA ACA AAG GAT AAT AAT TTT AAA CCA AAT GAT AAA AGT TTG		
lys gly lys lys tyr glu lys thr lys asp asn asn phe lys pro asn asp lys ser leu		
1321 / 441	1351 / 451	
TAT GAT GAG CAT ATT AAA AAA TAT AAA AAT GAT AAG CAG GTT AAT AAG GAA AAG GAA AAA		
tyr asp glu his ile lys lys tyr lys asn asp lys gln val asn lys glu lys glu lys		
1381 / 461	1411 / 471	
TTC ATA AAA TCA TTG TTT CAT ATA TTT GAC GGA GAC AAT GAA ATT TTA CAG ATC GTG GAT		
phe ile lys ser leu phe his ile phe asp gly asp asn glu ile leu gln ile val asp		
1441 / 481	1471 / 491	
GAG TTA TCT GAA GAT ATA ACT AAA TAT TTT ATG AAA CTA TAA AAG GTT ATA TAT		
glu leu ser glu asp ile thr lys tyr phe met lys leu		

FIGURE 9D

LSN.3'STOP -> 1-phase Translation

DNA sequence 1482 b.p. CAAGAACAACAA ... ATGAAACTATAA linear

1 / 1	31 / 11	
CAA GAA CAA CAA AGC GAT CTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA GAA		
gln glu gln gln ser asp leu glu gln gln glu arg ala lys glu lys leu gln glu gln		
61 / 21	91 / 31	
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT		
gln ser asp leu glu gln asp arg leu ala lys glu lys leu gln glu gln ser asp		
121 / 41	151 / 51	
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT CTA GAA CAA		
leu glu gln glu arg leu ala lys glu lys leu gln glu gln ser asp leu glu gln		
181 / 61	211 / 71	
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAG AGA CGT		
glu arg arg ala lys glu lys leu gln glu gln ser asp leu glu gln glu arg		
241 / 81	271 / 91	
GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA		
ala lys glu lys leu gln glu gln ser asp leu glu gln asp arg leu ala lys glu		
301 / 101	331 / 111	
AAG TTA CAA GAG CAG CAA AGC GAT TTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA		
lys leu gln glu gln gln ser asp leu glu gln glu arg ala lys glu lys leu gln		

FIGURE 10A

841 / 281	871 / 291	
GAA AAA ACA AAT AGA GAA TCT ATT ACA	ACA AAT GTT GAA GGA CGA AGG GAT ATA CAT AAA	
glu lys thr asn arg glu ser ile thr	thr asn val glu gly arg asp ile his lys	
901 / 301	931 / 311	
GGA CAT CTT GAA AAG AAA GAT GGT TCA ATA AAA CCA GAA CAA AAA GAT AAA TCT		
gly his leu glu lys lys asp gly ser ile lys pro glu gln lys glu asp lys ser		
961 / 321	991 / 331	
GCT GAC ATA CAA AAT CAT ACA TTA GAG ACA GTA AAT ATT TCT GAT GTT AAT GAT TTT CAA		
ala asp ile gln asn his thr leu glu thr val asn ile ser asp val asn asp phe gln		
1021 / 341	1051 / 351	
ATA AGT AAG TAT GAG GAT GAA ATA AGT GCT GAA TAT GAC GAT TCA TTA ATA GAT GAA GAA		
ile ser lys tyr glu asp glu ile ser ala glu tyr asp asp ser leu ile asp glu glu		
1081 / 361	1111 / 371	
GAA GAT GAT GAA GAC TTA GAC GAA TTT AAG CCT ATT GTG CAA TAT GAC AAT TTC CAA GAT		
glu asp asp glu asp leu asp glu phe lys pro ile val gln tyr asp asn phe gln asp		
1141 / 381	1171 / 391	
GAA GAA AAC ATA GGA ATT TAT AAA GAA CTA GAA GAT TTG ATA GAG AAA AAT GAA AAT TTA		
glu glu asn ile gly ile tyr lys glu leu glu asp leu ile glu lys asn glu asn leu		
1201 / 401	1231 / 411	
GAT GAT TTA GAT GAA GGA ATA GAA AAA TCA TCA GAA GAA TTA TCT GAA GAA AAA ATA AAA		
asp asp leu asp glu gly ile glu lys ser ser glu glu leu ser glu glu lys ile lys		
1261 / 421	1291 / 431	
AAA GGA AAG AAA TAT GAA AAA ACA AAG GAT AAT AAT TTT AAA CCA AAT GAT AAA AGT TTG		
lys gly lys lys tyr glu lys thr lys asp asn phe lys pro asn asp lys ser leu		

FIGURE 10C

1321 / 441	1351 / 451	
TAT GAT GAG CAT ATT AAA AAA TAT AAA AAT GAT AAG CAG GTT AAT AAG GAA AAG GAA AAA		
tyr asp glu his ile lys lys lys tyr lys asn asp lys gln val asn lys glu lys glu lys		
1381 / 461	1411 / 471	
TTC ATA AAA TCA TTG TTT CAT ATA TTT GAC GGA GAC AAT GAA ATT TTA CAG ATC GTG GAT		
phe ile lys ser leu phe his ile phe asp gly asp asn glu ile leu gln ile val asp		
1441 / 481	1471 / 491	
GAG TTA TCT GAA GAT ATA ACT AAA TAT TTT ATG AAA CTA TAA AAG GTT ATA TAT		
glu leu ser glu asp ile thr lys tyr phe met lys leu		

FIGURE 10D